A BIOLOGICALLY INSPIRED LEARNING TO GRASP SYSTEM

E. Oztop, M. A. Arbib
Department of Computer Science, University of Southern California, CA, USA

Abstract- **We propose and implement a learning to grasp system inspired from the development of reaching and grasping in infants, and the neurophysiology of the monkey premotor cortex. The system is composed of a virtual 19 DOF kinematics arm/hand and a learning mechanism that enables it to perform a successful grasp. The learning is based on "motor babbling". The model performs open hand reaches to the vicinity of the targets, which human infants younger than 4 moths of age appear to do. The contact of the hand with the object triggers an enclosure of the hand simulating the palmer reflex, characteristic to infants that are younger than 6 months of age. The varying degree of enclosure of each finger and the randomness in the reaching phase enables the system to explore the grasp configuration space. The learning scheme employed is a Hebbian one.**

***Keywords* - reaching, grasping, infant, Hebbian learning**

I. INTRODUCTION

We can perform a reach and grasp action for many objects in our daily lives effortlessly. However, the task is not trivial at all. The reach and grasp should be planned in ahead for the anticipation of the grasp configuration suitable for the object [1]. As humans have very dexterous hands the possible grasps that can be applied to an object is many. Iberall and Arbib [2] introduced the theory of *virtual fingers* and *opposition space.* The term *virtual finger* is used to describe the physical entity (one or more fingers, the palm of the hand, etc.) that is used in applying force and thus includes specification of the region to be brought in contact with the object (what we might call the "virtual fingertip"). Figure 1 shows three types of opposition: those for the precision grip, power grasp, and side opposition. Each of the grasp types is defined by specifying two virtual fingers, VF1 and VF2, and the regions on VF1 and VF2 which are to be brought into contact with the object to grasp it. Note that the "virtual fingertip" for VF1 in palm opposition is the surface of the palm, while that for VF2 in side opposition is the side of the index finger. The grasp defines two "opposition axes": the *opposition axis in the hand* joining the virtual finger regions to be opposed to each other, and the *opposition axis in the object* joining the regions where the virtual fingers contact the object. Visual perception provides *affordances* (different ways to grasp the object); once an affordance is selected, an appropriate opposition axis in the object can be determined. The task of motor control is to preshape the hand to form an opposition axis appropriate to the chosen affordance, and to so move the arm as to transport the hand to bring the hand and object axes into alignment. During the last stage of transport, the virtual fingers move down the opposition axis (the "enclose" phase) to grasp the object just as the hand reaches the appropriate position.

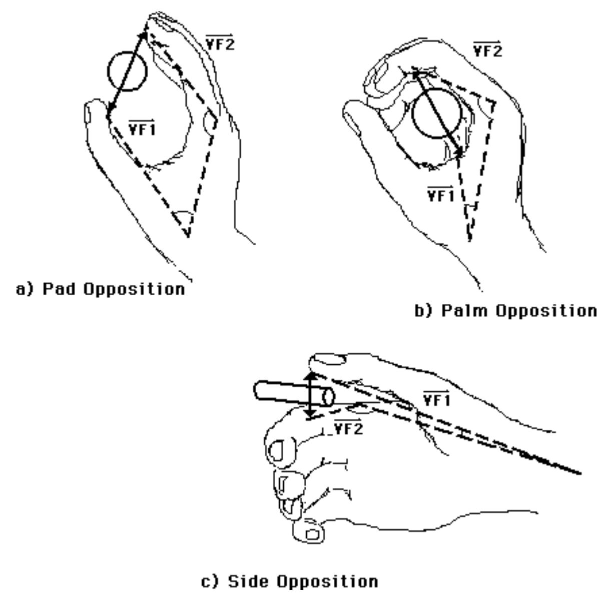

Figure 1. Each of the 3 grasp types here is defined by specifying two "virtual fingers", VF1 and VF2, which are groups of fingers or a part of the hand such as the palm which are brought to bear on either side of an object to grasp it. The specification of the virtual fingers includes specification of the region on each virtual finger to be brought in contact with the object. A successful grasp involves the alignment of two "opposition axes": the *opposition axis in the hand* joining the virtual finger regions to be opposed to each other, and the *opposition axis in the object* joining the regions where the virtual fingers contact the object. (adapted from [2])

The macaque inferior premotor cortex has been identified as being involved in reaching and grasping movements [3]. This region has been further partitioned into two sub-regions: F5, the rostral region, located along the arcuate and F4, the caudal part. The neurons in F4 appear to be primarily involved in the control of proximal movements [4], whereas the neurons of F5 are involved in distal control [3].

The onset of reaching and grasping marks a significant achievement in infants functional interactions with their surroundings. The advent of voluntary grasping of objects is preceded by several weeks in which infant engages in arm movements and fisted swipes in the presence of visible objects [5]. For many years, it has been accepted that the earliest accurate reaching behaviour is visually guided and appears around 3-5 months [6]. The term *visually guided reaching* generally refers to the infant's having available continuous vision of the hand and target, whereas *visually*

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25OCT2001	N/A	-

Title and Subtitle A Biologically Inspired Learning to Grasp System	**Contract Number**
	Grant Number
	Program Element Number
Author(s)	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Department of Computer Science, University of Southern California, CA	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes Papers from the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 25-28 OCT 2001, Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom.	
Abstract	
Subject Terms	
Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UU
Number of Pages 4	

elicited reaching refers to the vision of the target, followed by a ballistic hand movement. Clifton and her co-workers [6] questioned the visually guided reaching hypothesis. They tested seven infants repeatedly between 6 and 25 weeks of age to examine whether infants require vision of their hand when first beginning to reach for, contact and grasp objects. They used glowing or sounding objects for the dark condition. Infants first contacted the object in both dark and light conditions at almost the same ages (mean ages: for light, 12.3 weeks; for dark 11.9). Infants first grasped the object in light condition at 16.0 weeks and in the dark at 14.7 weeks (not a statistically significant difference). Clifton and her co-workers interpreted the results against the visual guidance hypothesis. They stated that since infants could not see their hand or arm in the dark, their early success in contacting the glowing and sounding objects indicated that proprioceptive cues, not sight of the limb guided their early reaching. Reaching in the light developed in parallel with reaching in the dark, suggesting that visual guidance of the hand is not necessary to achieve object contact either at the onset of successful reaching or in the succeeding weeks. It is also noteworthy to underline the fact that the infants showed great individual differences in onset behaviours. Onset touch varied between 7 and 16 weeks, while onset for grasp varied between 11 and 19 weeks. The greatest discrepancy (light versus dark conditions) in onset of reach and grasp was 4 weeks. There were three infants (out of seven) with this discrepancy. Interestingly for all the three infants the behaviour occurred earlier in the dark. However the findings does not conflict the traditional view of visual guidance for reaching that is reported in the literature as it would be unreasonable to claim that infants do not use vision information when it is available [6]. The infant, when contacted with the object, will occasionally try to grasp the object. The enclosure reflex will be with the infant until six moths of age and it will take 4 more weeks to stabilise the grasp [6]. However, the fractionated control of finger movements will not be possible since this task requires cortico-motoneuronal system, which has not been developed by the age of reflex grasping and early voluntary grasping [7]. Therefore, it is unlikely that the premotor specialisation for the different types of grasps that Rizzolatti group [3] has found be formed at this age yet. Infants will need to experience more to be able perform adult-like grasps. Before nine months old age, the infants grasp lack the anticipation of the orientation and the size of the object [8]; they adjust their grasps after touching the object. In contrast, the adults adjust their distance between the thumb and the other fingers according to the size of the object during the hand transport. Furthermore, infants younger than nine months old are physically able to vary their grip size, for they can spread their fingers farther apart once they have felt a large object [9].

II. METHODOLOGY

In this study the objective was to mimic the grasp development in infants and premotor functionality for grasp actions with a computer simulation. We developed the grasp learning system using Java language. The system is composed of a 19 DOF virtual arm that can be controlled manually through a user interface or automatically (e.g. for learning and testing) and a hybrid neural control circuit. We modeled the hand as 12 DOF (4 for the thumb and 2 for each finger). The wrist and shoulder have 3 DOF and the elbow has 1 DOF. We used forward kinematics to simulate the motion of the arm and hand. The system can detect the collisions of each segment on the arm with the objects in the workspace. Since in this study we focused on discovering grasp configurations appropriate for the objects, we did not include the learning of reach task (i.e. learning the inverse kinematics map). Instead we solved the inverse kinematics problem with the pseudo-inverse of the Jacobian of the forward kinematics transformation. The Figure 2 shows the virtual arm we used in our simulations after a precision grip. The neural part of the control represents the premotor area F5 of monkey. The circuit is trained using the feedback signaled by the attempted grasp action. The neural network we used informs the hand what level of enclosure is required for each hand joints. The conventional (i.e. non-neural) part of the controller performs the reach and orienting the hand towards the object. The learning we used is hebbian: the connections that are likely to be involved in producing successful grasp parameters are strengthened and the ones that tend to fail to do so are weakened.

Figure 2. A precision grip performed by the virtual arm model. The precision grip is generated using our non-neural grasp algorithm.

Before attempting to train the system, we implemented a conventional (i.e. non-neural) grasp (precision grip) solver. This solver planned the grasp shown in Figure 2. The algorithm we developed for the grasp planning is as follows.

- Determine the opposition axis to grasp the object.
- Compute the two (outer) points A and B at which the opposition axis intersects the object surface. They serve as the contact points for the virtual fingers that will be involved in the grasp.
- Assign the real fingers to virtual fingers. The particular heuristic we used in the experiments was the following.

If the object is on the right [left] with respect to the arm then thumb is assigned to the point A if A is on the left of [at a lower level than] B otherwise thumb is assigned to B. The index finger is assigned to the remaining point.

- Determine an approximate target position C, for the wrist. Mark the target for wrist on the line segment connecting the current position of the wrist and the target for thumb a fixed length (determined by the thumb length) away from the thumb target.
- Solve the inverse kinematics for only the wrist reach (ignore the hand).
- Solve the inverse kinematics for grasping. Using the sum of distance squares of the finger tips to the target contact points do a random hill climbing search to minimize the error. Note that the search starts with placing the wrist at point C. However, the wrist position is not included in the error term.
- The search stops when the simulator finds a configuration that makes the error close to zero (success) or after a fixed number of steps (failure to reach). In the success case the final configuration is returned as the solution for the inverse kinematics for the grasp. Otherwise failure-to-reach is returned.
- Execute the reach and grasp. At this point the simulator knows the desired target configuration in terms of joint angles. So what remains to be done is to perform the grasp in a realistic way (in terms of kinematics).
- The simplest way to perform the reach is to linearly change the joint angles from the initial configuration to the target configuration. But this does not produce a bell shaped velocity profile (not exactly a constant speed profile either because of the non-linearity in going from joint angles to end effector position.
- To get a bell shaped velocity we modify the idea of linearly changing the joint angles little bit. We simply modulate the change of time by replacing the time with a 3rd order polynomial that will match our constraints for time (starts at 0 climbs up to 1 monotonically). Note that we are still working in the joint space and our method may suffer from the non-linearity in transforming the joint angles to end effector coordinates. However our empirical studies showed that a satisfactory result, for our purposes, can be achieved with this strategy.

III. RESULTS

In this section we present the grasps configurations that our grasp learning system discovered. The training was performed as follows. The neural network representing area F5 of premotor cortex generates a (initially random) offset vector and a series of speed values for each joint of the fingers (initially random). The offset vector is added to the center of mass of the target object to obtain a reach target location. Note that this point may be in, on or outside the object. Then a reach is initiated to this point. The reach is performed with the palm open facing the object. During the transport the detection of a collision of causes a reflex enclosure of the hand. However, as mentioned earlier the speeds of the joint rotations are determined by the output of the neural network. If the enclosure leads to a successful grasp the connections that contributed to the generation of the parameters (offset and speed values) are strengthened. If the enclosure leads a failure then the connections that contributed to the generation of the parameters are weakened.

Figure 3 shows a learned power grasp directed to a sphere approximated as dodecahedron.

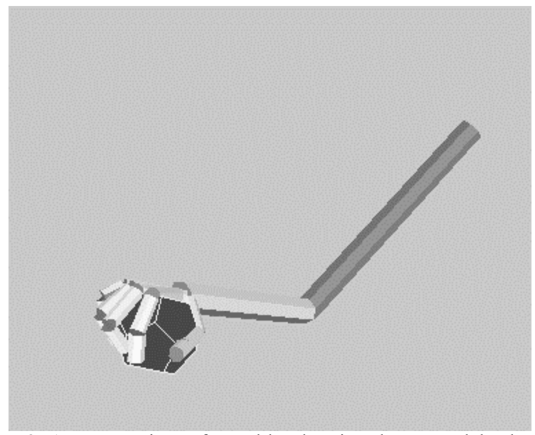

Figure 3. A power grip performed by the virtual arm model. The grasp parameters (hand offset and the joint speeds) are generated by the trained network.

Figure 4 demonstrates the discovery of palm opposition grasp (kind of power grasp without thumb being assigned to any virtual fingers) that has been introduced in Figure 1, part b. For this size object the network produced almost zero thumb speed. Where as for Figure 3, the thumb had to enclose the object so the thumb joints had non-zero speed.

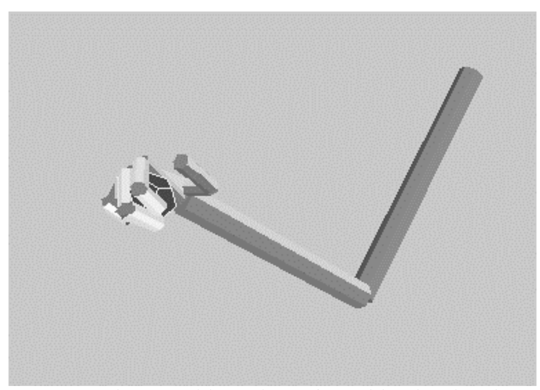

Figure 4. A palm opposition grip performed by the virtual arm model. The grasp parameters (hand offset and the joint speeds) are generated by the trained network.

The Figure 5 and 6 shows the precision grips that the network was able to learn. The network was not as good as in the earlier cases in generating this kind of grip. This is probably

because of the orientation of the palm is not learned. In order to produce a successful precision pinch it is not necessary that the palm normal coincides with the object. The further simulation will address the learning of wrist rotations as well.

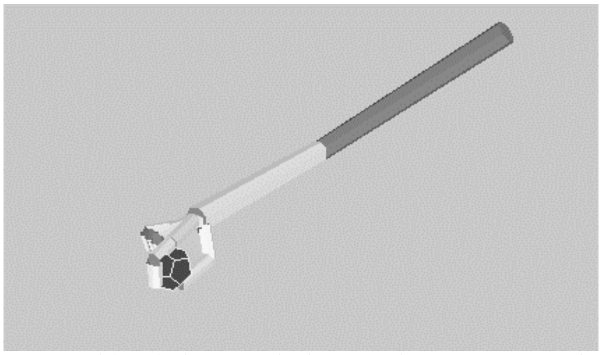

Figure 5. Precision pinch generated using network output parameters for the small sphere (approximated as a dodecahedron)

In figure 5 the discovered grasp is actually a mixture of power grasp (a single finger acting as virtual fingers 1 and 2) and precision grip. The network generated three virtual fingers for the grasp (two from the index finger and one from the thumb).

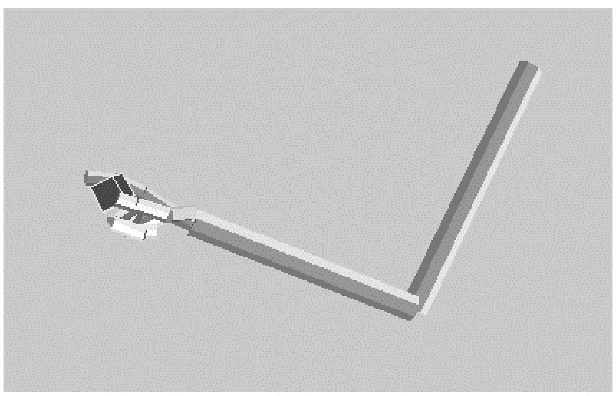

Figure 6. Precision grip generated using network output parameters for the cube shaped object.

However the precision grip in Figure 6 generated two virtual fingers (the thumb and index finger) which is more closer what usually the humans use for grasping small objects.

IV. Discussion and Conclusion

We have presented a hybrid system that can mimic grasp configuration learning by motor babbling. We showed that certain grasp configurations can be associated with certain objects with a simple mechanism such as palmer reflex that the infants born with. The palmer reflex enables the hand to enclose upond contact with object during a reach and the feedback on the success of grasp mediates learning during motor babbling. The shortcomings of our implementations are the kinematics (i.e. no dynamics) implementation of the arm/hand apparatus and the the lack of detailed modeling required to transfer the haptic and proprioceptive feedback from the hand to the F5 via somatosensory cortex which is the current focus of our study. It would be very interesting to implement a very accurate 3D model of the hand to see whether we can produce the daily life grasping examples. Our simulation system has not force simulation, however in reality considerable amount of grasp planning is devoted to weight anticipation and balancing the torque generated by the gravity. Nevertheless our grasp learning system is a step towards a full dynamics simulation with a full neural implementation, which can discover realistic grasps.

Acknowledgment

This work was supported in part by a Human Frontier Science Program grant to Arbib which made possible extensive discussions of data on the premotor cortex and monkey grasping circuit with Giacomo Rizzolatti, Vittorio Gallese, to whom we express our gratitude.

References

[1] M. Jeannerod, M.A. Arbib, G. Rizzolatti, H. Sakata, “Grasping objects: the cortical mechanisms of visuomotor transformation,” *Trends Neurosci.,* 18, 314-320, 1995.

[2] T. Iberall, M.A. Arbib M.A. “Schemas for the Control of Hand Movements: An Essay on Cortical Localization.” In *Vision and action: the control of grasping,* Goodale M.A., editor. Norwood, NJ: Ablex, 163-180, 1990.

[3] G. Rizzolatti, R. Camarda, L. Fogassi, M. Gentilucci, G. Luppino, M. Matelli. “Functional Organization of Inferior Area 6 in the Macaque Monkey II. Area F5 and the Control of Distal Movements,” *Experimental Brain Research,* 71: 491-507, 1988.

[4] M. Gentilucci, L. Fogassi, G. Luppino, M. Matelli, R. Camarda, G. Rizzolatti, “Functional Organization of Inferior Area 6 in the Macaque Monkey I. Somatotopy and the Control of Proximal Movements,” *Experimental Brain Research,* 71: 475-490, 1988.

[5] C. von Hofsten, “Developmental changes in the organization of prereaching movements,” *Developmental Psychology*, 20:378-388, 1989.

[6] R.K. Clifton, W.M. Darwin, D.H. Ashmead, M.G. Clarkson, “Is Visually Guided Reaching in Early Infancy a Myth? ” *Child Development*, 64: 1099-1110, 1993.

[7] C. Lantz, K. Melen, H. Forssberg, “Early infant grasping involves radial fingers,” Developmental Medicine and Child Neurology, 38: 668-674, 1996.

[8] D. A. Rosenbaum, Human motor control. Academic Press. Amherst, Massachusetts, 1991.

[9] C. von Hofsten, L. Ronnqvist, “Preparation for grasping an object: A developmental study,” *Journal of Experimental Psychology:Human Perception and Performance,* 14:610-621, 1988.